

Le rôle de la répétition dans la représentation du sens et son approche statistique par la méthode "ALCESTE"¹

MAX REINERT

Noms et sons n'ont-ils pas été données aux choses pour que l'homme y prenne plaisir? C'est une douce folie que le langage: grâce à lui l'homme passe en dansant sur toutes les choses.

Nietzsche²

Résumé

Dans cette étude, l'auteur cherche à formuler ce qui lui apparaît fonder sémiotiquement sa pratique d'analyse statistique de données textuelles. Après une courte introduction autour de la question: 'Quel peut être le rôle de l'approche statistique dans l'analyse du sens d'un discours, et quelle en est la limite', la première partie aborde, à partir de la sémiotique Peircienne, le rôle de la répétition dans la constitution du sens. Dans la seconde, l'auteur applique cette conception à la méthode statistique d'analyse de discours ALCESTE [sigle pour 'Analyse des Lexèmes Cooccurrents dans un Ensemble de Segments de Textes']. En conclusion, il souligne l'impossibilité d'un 'point de vue' externe pour définir le sens. On trouvera également quelques résultats et commentaires succincts à propos de l'analyse 'ALCESTE' du texte de cet article en annexe.

Introduction

L'objectif que nous nous proposons est d'aborder la notion de signe de manière très générale en partant de la notion de répétition, afin de suggérer que l'approche du sens³ des discours dans certains de ses aspects ne présuppose pas nécessairement l'utilisation de modèles d'analyse de la signification linguistique.

Les méthodes statistiques d'analyse des données introduites par J. P. Benzécri⁴ dès les années 1970 ont montré leur efficacité dans le domaine des analyses textuelles (Lebart et Salem 1994) . . . Du moins, c'est ce que les

utilisateurs de ces méthodes pensent. Cela dit *quel peut être le rôle de l'approche statistique dans l'analyse de discours*,⁵ et *quelle en est la limite*? Cet essai de clarification se fera autour de notre expérience particulière de la méthode *ALCESTE*.

Commençons par quelques questions: Qu'est-ce qu'une telle méthode statistique peut *représenter* d'un discours? Peut-on parler *du contenu* d'un discours sans analyser sa signification linguistique? N'y a-t-il pas contradiction à proposer une méthode quasi-automatique d'analyse⁶ alors même que le sens d'un discours ne peut se constituer en dehors de l'engagement du locuteur?⁷

Pour le sociolinguiste Achard (1993), le discours se définit par '*l'usage du langage en situation pratique, envisagé comme acte effectif, et en relation avec l'ensemble des actes (langagiers ou non) dont il fait partie*'. Dans le cadre de cette conception, l'analyse de discours ne peut donc être tentée en dehors de l'engagement de l'analyste, en tant que *sujet*, car c'est par cet engagement que le discours peut prendre un sens (Achard 1997). Cela dit — et nous montrerons que ce n'est pas contradictoire — cette singularité du discours, en tant qu'acte, est également assujéti à la *répétition*. Dans *l'ordre du discours*, Foucault écrit à propos du commentaire ce que nous assumons pour tout discours: '*Il doit, selon un paradoxe qu'il déplace toujours mais auquel il n'échappe jamais, dire pour la première fois ce qui cependant avait été déjà dit et répéter inlassablement ce qui pourtant n'avait jamais été dit. Le moutonnement indéfini des commentaires est travaillé de l'intérieur par le rêve d'une répétition masquée*'. Foucault ajoute: '*Le nouveau n'est pas dans ce qui est dit, mais dans l'événement de son retour*'.

Répétition et sémiose

La notion de *répétition* est sans doute trop fondamentale pour sortir du langage courant. Certes, à la suite de Freud, elle est promue par Lacan (1964) comme un des '*quatre concepts fondamentaux de la psychanalyse*'.⁸ Avant Freud, Kierkegaard lui consacra un ouvrage; Cela n'a pas suffi pour que cette notion apparaisse comme entrée du '*Vocabulaire technique et critique de la philosophie*' de Lalande. Quant aux grammairiens, ils n'en parlent que pour l'éviter (l'anaphore). Dans le *Lexis*, le lemme 'répétition' est introduit comme nominalisation de 'répéter' avec cette simple définition: '*Retour de la même idée, du même mot, du même fait*'. Définition succincte, certes, mais dont la résonance nous a suggéré une affinité particulière avec les trois modalités du *signe Peircien*:

- *L'icône*, pour l'idée, en tant qu'elle peut être l'idée vague d'un *quelque chose*;

- *L'indice*, pour le fait en tant qu'il est circonscrit à une *expérience réelle d'un objet*;
- Et le *symbole*, pour le mot, en tant qu'il est lisible, déchiffirable par convention de représentation.

La rencontre entre cette définition de la répétition et des trois modalités du signe Peircien n'est pas fortuite, pour nous. Elle marque la place centrale de la répétition dans la définition du signe, et plus précisément dans la mise en rapport d'un signe à son objet.

Pour approfondir cette idée encore vague remarquons tout d'abord que pour Peirce, un signe ne se définit pas en soi mais dans le mouvement d'une sémiose. Il naît d'une apparence (*representamen*) pour quelqu'un. Mais cette sollicitation, par l'intérêt qu'elle exprime, doit engager l'interprète vers autre chose, vers un *objet* que le *representamen* cache et révèle à la fois. Donc un signe renvoie à un *objet*, mais cet *objet* ne peut être interprété que par un autre signe, que Peirce appelle son *interprétant*. De cet *objet* du signe dont parle Peirce, disons seulement qu'il est ce qui circule d'un signe à l'autre, qu'il est *ce qui se répète* pour quelqu'un, qu'il est surtout le '*même objet*' pour l'*interprète* du signe. Peirce ne parle d'interprète que '*parce qu' (il) désespère de faire comprendre (sa) propre conception plus large*'.⁹ Pourtant, il ne cesse pas d'y faire référence. C'est le premier point problématique de la théorie Peircienne: Quelle théorie du *sujet* convoque-t-elle? Peirce n'emploie d'ailleurs le terme de '*sujet*' que pour l'opposer à celui de '*prédicat*'. On constate cependant son influence sur le courant psychanalytique et plus particulièrement Lacanien.¹⁰ Quoi qu'il en soit, sa définition du signe rappelle singulièrement la définition du *signifiant* Lacanien par son côté circulaire — coté circulaire qui a également été souligné par Benveniste (1974: 44–45), mais pour le critiquer:¹¹

Peirce: un signe (representamen) représente son objet pour un autre signe (son interprétant)

*Lacan (1973: 222): un signifiant représente le sujet pour un autre signifiant.*¹²

Entre l'*objet* Peircien et le *sujet* Lacanien, il y a comme un jeu de miroir qu'évoque cette remarque de Bachelard (1949) en introduction de sa '*Psychanalyse du feu*': '*Il suffit que nous parlions d'un objet pour nous croire objectifs. Mais par notre premier choix, l'objet nous désigne plus que nous ne le désignons et ce que nous croyons nos pensées fondamentales sur le monde sont souvent des confidences sur la jeunesse de notre esprit*'.

Notre point de départ est le suivant: Dans la sémiose, c'est parce qu'il y a glissade d'un signe à l'autre, en acte, que le *sujet* se constitue dans sa visée d'un *même objet*. Plus précisément, on ne peut attribuer au '*même*' de la *répétition*, ni la permanence d'une chose, ni l'invariabilité d'un point de

vue. C'est le lieu où, justement, semble s'organiser dualement un *monde* pour un *sujet* et un *sujet* pour un *monde*. Mais cette dualité boîte; Elle est en porte-à-faux par cet écart, générateur d'un temps propre par lequel le sujet existe, entre ce qui s'offre à lui et ce qu'il croit saisir, et dont le signe marque justement le malentendu.

Si la *répétition* est bien dans ce qui noue un signe à sa temporalité, pour un sujet, l'*objet* de ce signe est donc d'abord, *objet de valeur*, dans la mesure où il semble circuler d'un signe à l'autre. Dans cette perspective, le sens d'un signe est ce mouvement qui va d'un signe à l'autre, *mouvement basé sur la perte de l'objet*, puisque tout signe, en devenant tangible, prend la place de son *objet* en se montrant *autre*.

Ce mouvement des signes n'est cependant pas gratuit. Il doit être compatible avec des exigences réelles, liées à une activité et à l'existence de contraintes réelles, supportées par des *sujets* en situation et sans la dynamique desquels rien ne peut faire signe.

En définitive, ce sont les activités qui structurent le sens, si l'on entend par là, que la circulation des signes dans ces activités aura tendance à se stabiliser dans des réseaux d'échange, bien avant que le sens de cette circulation active soit représenté. C'est notre principale hypothèse de travail pour l'analyse du sens. *Cela aboutira à chercher à modéliser non pas la signification d'un texte mais sa trace comme activité.*

Revenons à Peirce. Sa définition du signe la plus citée est la suivante:

Un signe, ou representamen, est quelque chose qui tient lieu pour quelqu'un de quelque chose sous quelque rapport ou à quelque titre. Il s'adresse à quelqu'un, c'est-à-dire crée dans l'esprit de cette personne un signe équivalent ou peut-être un signe plus développé. Ce signe qu'il crée, je l'appelle l'interprétant du premier signe. Ce signe tient lieu de quelque chose: de son objet. Il tient lieu de cet objet, non sous tous rapports, mais par référence à une sorte d'idée que j'ai appelée quelquefois le fondement du representamen. (O.c., p. 121)¹³

Commentaire: a) Un signe tient lieu pour quelqu'un de quelque chose; b) Mais ce *quelque chose* ne peut être appréhendé que par un autre signe, son interprétant; c) Cet interprétant n'est pas quelconque, il doit être consistant par rapport à toutes les sémoses précédentes. C'est en tout cas, notre interprétation de '*il tient lieu de cet objet . . . par référence à une sorte d'idée*' . . . qui était déjà dans ce qui fonde le premier signe . . . du fait des sémoses antérieures. Dans l'idée de Peirce, il y a sans doute du logique derrière cela. Il y a également, selon nous, une dépendance liée à un parcours singulier; ce qu'on peut appeler une *logique de sujet*.¹⁴

Partant de cette définition, Peirce distingue ensuite trois modes de fonctionnement d'un signe relativement à son *objet*: *l'icône, l'indice et le*

symbole. Très succinctement, *l'icône* évoque son *objet* en vertu de sa nature interne; *l'indice* 'est réellement déterminé' par son *objet* comme, par exemple, la fumée par rapport au feu; *le symbole* n'est relié à son *objet* que par une convention ou une loi d'interprétation, comme le feu rouge pour le code de la route.

Comme nous l'avons déjà signalé, nous soupçonnons sous ces trois modes de fonctionnement du signe relativement à son *objet*, trois formes de *répétition* . . . qui, en définitive, fondent *l'objet* aussi bien que le *sujet*. Dans cette perspective, *l'objet* du signe est davantage *objet de valeur*; c'est-à-dire, ce que poursuit un *sujet*. Mais ce que poursuit un *sujet* est également dans ce qui le pousse à agir. *L'objet* recherché, étant à la source du désir de recherche, est par conséquent *dynamique*. Il s'impose du fait de l'usage actif et répété des signes. Ce sera là notre interprétation de la notion d'*objet dynamique* chez Peirce.¹⁵

Une première forme de répétition, l'icône

Notre utilisation de la notion d'*icône* est peut-être différente de celle de certains Peirciens. Nous ne prétendons pas non plus être fidèle à Peirce lui-même, ne serait-ce que parce que nous introduisons la notion d'*icône* en relation avec la notion de *répétition* . . . ce qui conduit à envisager le signe en relation avec un *sujet* pour qui ces *répétitions* prennent sens. Cela dit, cette approche ne nous semble pas contradictoire avec l'approche Peircienne. On partira de cette définition de 1904: '*Je définis une Icône comme un signe qui est déterminé par son objet dynamique en vertu de sa nature interne (. . .)*' (O.c., p. 32).¹⁶ Certes, ce '*en vertu de sa nature interne*' laisse un peu rêveur. Prenons cette autre définition de Peirce: '*Une icône est un signe qui posséderait le caractère qui le rend signifiant, même si son objet n'existait pas*' (O.c., p. 139).¹⁷ Elle semble montrer qu'une *icône* ne peut être définie simplement comme un signe *ressemblant* à son *objet*, sinon Peirce le dirait . . . Il y aurait d'ailleurs une contradiction à l'affirmer.¹⁸ Pour notre part, nous définirons l'*icône* en rapport avec la *répétition*, par l'effet associatif qu'elle provoque chez un *sujet*. L'association en tant que résonance particulière touche à un *aspect physique* voire *musical* de la *ressemblance*. Un signe se détache d'un fond comme une note musicale se détache d'un fond sonore quand elle entre en résonance avec le diapason. Subjectivement cela peut être perçu sous la forme d'une émotion particulière en diapason avec la *répétition*. Elle a un *contenu émotionnel immédiat* (priméité). Ce qui fonde la *transparence* du signe iconique est justement ce qui l'ouvre à un *contenu* immédiat.

Cela conduit à préciser notre notion de *contenu*: *Un contenu est introduit par un signe*, non pas en tant que *re-présentation médiate*, mais en tant que *contenu immédiat*, effet d'une *résonance* particulière. Ainsi, pour nous, un *contenu* n'est ni un concept, ni un *signifié*. *En tant qu'il est vécu comme acte, il est l'expression immédiate de l'aspect iconique du signe*.

Ainsi, *comme forme première de répétition, l'icône* se distingue d'un fond par sa saillance.¹⁹ L'immédiateté du *contenu* recouvre la mémoire d'un parcours de vie. Et *l'icône*, en se dévoilant, donne à ce *contenu* le premier caractère sensible qui l'externalise;²⁰ elle éveille la *conscience d'un quelque chose*. Elle est perçue par l'interprète comme pure transparence, comme le débordement d'un *déjà connu*; Mais ce trop plein émotionnel de l'*objet* n'exprime pas sa saisie conceptuelle. Ainsi *l'icône* évoque un *même objet* qui ne cesse de se voiler et de se dévoiler, à la fois fuyant et fascinant, un et multiple, et dont les échos se prolongent sans pourtant lui donner une consistance stable. La fascination de *l'icône* tient à cette distinction particulière liée à la résurgence d'un passé expérientiel qu'elle voile pourtant. Prendre *conscience d'un quelque chose* ne peut commencer qu'après avoir perçu une icône.²¹

Mais dans son *fondement, l'icône* est d'abord vibration plus que forme, miroitement récurrent d'un '*même*' inaccessible, idée vague et pourtant pénétrante, émotion plus que perception. Dans nombre de mythes, l'homme naît de la glaise. *Matière* ou *contenu*, n'est-ce pas ce qui fourmille comme la multiplication d'un *même objet voilé*, bien avant de prendre forme? Bachelard (1942) a été particulièrement sensible à cet appel *iconique* des quatre *matières élémentaires*, l'eau, la terre, l'air et le feu, qu'il considère comme des '*hormones de l'imagination*'.

Plus spécifiquement au niveau du discours poétique, ce *processus d'iconification* de la *répétition* peut se conjuguer avec des phénomènes comme l'allitération (aspect phonétique) ou comme *l'isotopie* (aspect sémantique). Ce retour du *même* est multiplication d'un *même*. Il est écho. L'insistance répétée de son retour remédie à l'instabilité de la forme. C'est par le prolongement de cet appel répété que le *sujet* prend la mesure de cet autre lui-même qui est son *monde* propre. Ainsi dans son propre mouvement pour saisir le *monde*, il se saisit lui-même. Nous retrouvons trace de ce processus dans la création artistique. Le style,²² c'est ce qui se donne à profusion à la fois *monde* et *sujet*: Cézanne, Braque, Matisse.

Notre principale hypothèse de travail est que ce type de répétition iconique peut être repéré à un niveau très archaïque dans tout énoncé par la simple cooccurrence des mots pleins, par l'effet de résonance que le contenu introduit entre eux (aspect associatif).

Prenons l'exemple de ce vers de Reverdy, extrait du poème '*Courte vie*' du recueil '*Les sources du vent*': '*L'heure pleine est passée sur une autre qui*

sonne. *Les pas des voyageurs courent déjà plus loin* ; La simple présence des mots pleins comme *heure, passée, pas, voyageurs, courent, loin* dessine dans ce cas une isotopie du passage indépendamment de la manière dont sont reliés ces mots au niveau syntaxique.

Cette première impression isotopique de l'énoncé, que nous appellerons par référence à Peirce le *fondement topique* de l'énoncé, est ce qui fonde à la fois le *sujet énonciateur* et l'*objet de l'énonciation*, non comme *forme*, mais comme *appel*. Par ce *fondement*, l'icône apparaît donc comme le trait d'union entre un passé qui se dérobe — et pourtant revient — et un présent qui se démultiplie — et pourtant se cache — comme l'image démultipliée des miroirs, démultiplication appréhendée par les sens, d'abord comme matière ou comme *contenu*, avant de s'effacer dans la transparence d'une forme.

Remarque.

- a. Au plan de la méthodologie, contrairement à la recherche d'*isotopie*,²³ nous ne chercherons pas à thématiser le *fondement topique* d'un énoncé ou à l'identifier par un *sème*. *Le fondement topique doit garder l'indétermination d'une matière que le sème n'a pas*. Il n'est véritablement qu'un *fond d'émergence* de l'*objet* pour un regard possible. De ce point de vue, nous sommes en accord avec Bachelard: la forme ne s'oppose pas au fond, mais la forme se développe à partir d'un fond qui est d'abord matière, et elle se développe en interaction avec un regard, le regard d'un *sujet* naissant à sa propre conscience.
- b. Au plan de la méthode *ALCESTE*, notre hypothèse sur la notion de 'plénitude' des *mots pleins* est davantage liée à cette forme archaïque de *répétition iconique* chez le *sujet*. Un *mot plein* est l'*écho d'une icône* dont le *contenu* est lié aux expériences passées du *sujet*. Il porte en lui la marque des lieux où il fut reconnu. Pour le *sujet*, reconnaître une *icône* est une manière de se remémorer en tant qu'elle exprime un attrait particulier. Par la notion de *fondement topique*, on désire seulement souligner la tendance des *mots pleins* d'un énoncé à connoter une même *icône*, une même matière ou un même *contenu*.
- c. Au plan notionnel, on doit donc séparer:
 - La notion de *plénitude* d'un *mot plein* qui renvoie à la notion d'*icône*: la valeur attribuée au mot vient de la *répétition* liée à l'expérience singulière d'un *sujet*;
 - De la notion de *signifié*, qui renvoie au *système* de la langue: la valeur attribuée vient de sa distribution sur les *formes lexicales* plus ou moins normalisées.

Il est cependant logique que le sens commun de ces deux mots ait tendance à se confondre du fait d'un ajustement constant des usages et des règles d'une langue dans une société donnée.

Une deuxième forme de répétition, l'indice

On l'a compris, l'*icône* est perçue par ce mouvement du *sujet* à la recherche d'un *objet* qu'il croit saisir un instant, effet de *résonance* imaginaire propre à l'*icône*. L'*indice*, au sens Peircien, évoque au contraire un existant, hic et nunc, dont on ne parle pas directement mais qui s'impose justement à travers son *indice*. Plus précisément, pour Peirce, '*un indice est un signe qui renvoie à l'objet qu'il dénote parce qu'il est réellement affecté par cet objet*' (O.c., p. 140).²⁴ Par exemple, la fumée peut être un indice du feu ou la position de la girouette, un indice de la direction du vent. Cela dit, un *indice* doit d'abord être perçue comme *icône*; Par exemple, la fumée; Elle prend ensuite un autre sens comme *indice*, pour l'interprète, du fait de la pression de la situation. Ce dédoublement du sens dans le signe est une conséquence de sa valeur *d'indice*.

Nous partageons la conception d'un *Réel* non représentable par opposition à la *Réalité* qui est de l'ordre de la représentation ou de la mimesis. Il est indubitable cependant qu'on ne puisse l'ignorer (le réel). L'*indice* est justement ce dont on doit tenir compte pour cela. Et la *répétition* est là pour nous rappeler que '*le réel est ici ce qui revient à la même place*' (Lacan 1964: 59).

Par rapport à l'attente du *sujet*, l'*indice*, comme *répétition*, s'impose plutôt négativement comme sanction. Quelque chose résiste et donne un son fêlé à la belle *résonance iconique*. Mais ce *quelque chose* reste masqué. Il ne se révèle que dans le sentiment d'étrangeté de l'*icône*. Ainsi l'*icône*, en devenant *indice*, finit par s'effacer comme *icône*. Reste une absence, un arrêt, une ponctuation. Le *réel* se fait trou, vertige. L'*indice* aurait donc comme limite ultime une béance. L'évanouissement de l'*icône* souligne la *coupure* qui en marque la fin. Selon cette conception, prendre conscience d'un *indice*, c'est vivre l'inquiétude du non reconnu, vivre la rencontre avec le *réel*, *rencontre ratée* cependant puisque l'*indice* le voile et ne le révèle que dans l'évanouissement de l'*icône* qui le soutient.

Au plan du discours, si par la notion de *fondement topique*, on ancre l'énoncé dans un *imaginaire*, une *matière*, un *lieu origine*, un *contenu*, celui-ci se termine par un blanc, une coupure, qui en marque le dessaisissement. En cela, la coupure met en valeur le non-dit indicial de l'énoncé. Elle rend compte rétrospectivement d'un *autre dit*, d'un 'inter-dit', pour reprendre la formule suggestive de Lacan.

L'*indicialité* de l'énoncé implique, pour celui-ci, une adresse du fait de la référence à une autre voix. Elle exprime la présence de l'autre dans le discours. Elle introduit au dialogisme. Pour Bakhtine (1977), par exemple, un énoncé se définit davantage comme réponse à un autre énoncé que par sa structure interne. L'énoncé devient *polyphonique*.²⁵ Mais dans son

adresse à un *autre*, au moment même où il se dédouble, l'énonciateur finit par se voir lui-même. L'indicialité se fait indexicalité.²⁶ Il peut également lui sembler que ce soit un *autre* qui parle dans son propre énoncé. L'énoncé s'affirme comme hétérogène par constitution, pour reprendre le terme de J. Authier-Revuz (1982). En effet, il entremêle si intimement cette autre voix à ce qui est dit que l'on ne peut guère espérer en démêler le sens autrement que dans le parti pris d'une nouvelle lecture.

Au plan méthodologique, si chaque rupture dans un discours est l'expression, en négatif, de l'émergence d'une autre parole, elle ouvre également, positivement, à une nouvelle prise de parole, à un nouvel énoncé. Et la multiplicité des énoncés d'un discours, leur relatif emboîtement, est une trace fondamentale de cette dynamique polyphonique des énonciateurs, éventuellement conflictuelle. Le discours porte une trace intime de cette dynamique énonciative. Il la porte dans son nom²⁷ même qui articule 'dis' et 'cours', 'dis' exprimant la discontinuité, ce qui s'impose sans pouvoir vraiment se dire, et 'cours', la fluidité d'un cours d'eau, ce qui se dit et se répète inlassablement sans pouvoir vraiment exister. Son nom scelle ainsi, par ses deux syllabes, l'être de tout discours qui naît de ce qu'il n'est pas, de l'hétérogénéité qui le constitue.

Ainsi le discours est une suite de ruptures et d'échos d'amplitudes diverses. Il commence dès que l'on associe une consonne (dis) à une voyelle (cours), se prolonge avec la construction des syllabes jusqu'à l'organisation des parties d'un discours. Par exemple, dans une proposition, le sujet (dis) s'efface pour faire place au prédicat (cours). Dans un énoncé, le thème (dis) glisse d'un énoncé à l'autre, sans vraiment pouvoir se dire, pour laisser place au rhème (cours), etc. Dans ce jeu de cache-cache des ruptures et des échos, il est bien difficile de saisir objectivement où commence et où se termine un énoncé.

L'idée sous-jacente à cette énumération est que toute énonciation se développe comme une fractale.²⁸ Elle mêle, en une même matérialité du discours, tout un jeu de ruptures et d'échos à des échelles rythmiques très diverses. Ce jeu multiple semble cependant basé sur un même processus commun récursif dont le prototype est formalisé par la proposition logique, et qui consiste à *joindre* un *indice* (son sujet) avec une *icône* (son prédicat) par un *symbole*, comme on le verra. Soulignons cependant une double anomalie par rapport à notre présentation. D'abord dans une proposition logique, l'ordre de la sémiose y semble inversé. En effet, et contrairement à notre présentation, le sujet venant temporellement avant le prédicat, l'*indice* semble devancer l'*icône*. *C'est que la proposition ne recouvre pas une sémiose, mais un changement de sémiose*. Plus précisément, la coupure temporelle bordant la fin de l'*indice* dans la sémiose est la place que va occuper la *copule* dans la proposition. *Ainsi cette coupure devient*

lien. C'est par elle qu'on va pouvoir accéder à une nouvelle *icône*, qui est le *contenu* de la proposition, et qui, dans une proposition logique, se confond avec le prédicat. Pour comprendre ce renversement de l'ordre de la sémiologie et la *prise de conscience* particulière qui l'accompagne (le jugement), il faut d'abord bien comprendre en quoi une rupture (la *coupure indicielle*) peut faire lien (la copule). C'est cela justement qui définit pour nous le *symbole*.

Une troisième forme de répétition: le symbole

Pour Peirce, 'un symbole est un representamen dont le caractère représentatif consiste précisément en ce qu'il est une règle qui déterminera son interprétant' (O.c., p. 161).²⁹ Cette définition reste ambiguë pour plusieurs raisons et déjà par le fait que la règle peut être explicitement représentée, ou bien, elle peut être implicite, active à travers un usage, des *habitus*.³⁰ Cette définition s'interprète cependant clairement dans le cas d'un signe dans une théorie mathématique. Le symbole représente une règle. Il n'est donc pas un signe matériel comme l'*icône* ou un signe dynamique comme l'*indice*, bien qu'on ne puisse le représenter que par de tels signes; C'est un signe formel dont on distingue le type de l'occurrence.³¹ Mais dire qu'un symbole est un signe formel n'explique rien. Il faut encore remonter à la source de cette conception pour retrouver dans le symbole, ce qui émerge comme limite d'une série, d'une répétition, d'un usage.

Selon l'*Encyclopaedia Universalis*,³² le sens étymologique du mot 'symbole' est dérivé du verbe 'je joins' en grec. Il se définit à partir d'un objet partagé en deux. Par exemple, la possession partagée par deux comploteurs des deux morceaux d'un même billet déchiré leur permet de se reconnaître sans se connaître. Ainsi, un symbole exprime étymologiquement une déchirure qui fait lien. C'est justement le sens principal que nous lui donnons: une coupure qui fait lien . . . avec cependant cette condition, qu'il fait lien pour quelqu'un; Et, à cause de cela, précédé par l'usage de ce lien, bien avant d'avoir une conception de la règle.

Ce qui se sépare était lié et le symbole exprime cette réversibilité de l'acte dans la représentation. Par exemple, l'acte même de la déchirure du billet se réfléchit dans l'acte qui réunit les deux morceaux. Le tombé-juste³³ de cet acte en retour constitue un nouveau signe dont le sens est symbolique.

Contrairement à l'*icône*, qui se définit par son contenu, le symbole est introduit par un acte dont l'effet est prévisible. Il se pose donc comme une sorte de calcul. C'est en tant que le geste de recollement tombe juste, comme on l'attend, que le symbole est apprécié dans ce qu'on appelle un jugement. Donc le symbole est d'abord vécu à travers des gestes, qui, de se répéter, semble tomber juste, en acte, pour leur promoteur. S'il est

représenté par une marque, le symbole est sans rapport direct avec un contenu, mais l'usage d'un symbole peut conduire à le doter du contenu vécu en situation dans la mesure où celle-ci a tendance à se répéter, à faire écho, résonance. Il redevient alors icône. Cela dit, la spécificité du symbole est de rendre visible son propre geste par la réversibilité qu'il connote. Cette possibilité de présentifier l'acte passé conduit à nier l'impossibilité d'un retour dans le temps. C'est en cela que le symbole ouvre à l'espace des re-présentations.

Cette définition du symbole rejoint-elle la définition du signifiant chez Lacan et plus précisément, du trait unaire, par lequel, un acte devient objet de représentation? Pour cet auteur, le trait unaire pourrait avoir été conçu originellement par le chasseur comme marque, entaille, coche de l'animal tué:

Pour l'illustrer je vous rappellerai que la chose peut se présenter de la façon la plus simple dans le trait unaire. Le premier signifiant, c'est la coche, par où il est marqué, par exemple, que le sujet a tué une bête, moyennant quoi, il ne s'embrouillera pas dans sa mémoire quand il en aura tué dix autres. Il n'aura pas à se souvenir de laquelle est laquelle, et c'est à partir de ce trait unaire qu'il les comptera. (Lacan 1973: 159)

Que dire, en quelques mots, à propos du lien entre la notion de symbole chez Peirce et la notion de signifiant chez Lacan? Nous partirons de la proposition de M. Balat (2000: 17–36) qui propose une lecture du signifiant Lacanien comme *representamen*. Rappelons que pour Peirce, 'un *representamen* est le sujet d'une relation triadique, avec un second appelé son objet, pour un troisième appelé son interprétant' (O.c., p. 117 [1.541]). La notion de sujet est, au sens Peircien, ce qui s'impose, ici et maintenant, en situation.³⁴ Cette imposition du *representamen* signifie déjà que celui-ci convoque quelque chose chez quelqu'un dans sa capacité d'être distingué, séparé, capacité qui est la mémoire d'une histoire que celle-ci soit individuelle (les habitudes), sociale (les *habitus*) voire génétique (la couleur par exemple). Cette capacité du *representamen* de faire signe, d'être saillant dans une situation, est essentielle, car c'est par elle que de nouveaux contenus peuvent entrer en association, en résonance, avec le vécu des situations où ce *representamen* a été reconnu. L'habitude établie, cette forme de reconnaissance du signe ne convoque plus la conscience. Autrement dit, l'intérêt de l'interprète n'est plus dans le *representamen*, mais dans ce qu'il annonce de la situation, par expérience. Ainsi l'aspect formel passe au second plan pour lui. Ce qui est formellement reconnu devient transparent pour l'interprète dont l'intérêt se focalise sur le contenu; C'est-à-dire, sur la capacité du *representamen* à animer une nouvelle icône.

Selon cette conception, le signifiant est cette place aveugle du symbole dans le representamen qui contraint l'icône qu'il convoque chez un sujet à certains contenus. Ce n'est pas contradictoire avec l'interprétation de Peirce même si elle ne la recouvre pas exactement: '(. .) le representamen détermine son interprétant à entretenir la même relation triadique avec le même objet pour quelque interprétant' (O.c., p. 117 [1.541]).

Le signifiant ainsi présenté se rapproche de son homologue Saussurien par le fait qu'il introduit un contenu, mais s'en distingue, comme le signifiant lacanien, par la fluidité du signifié (M. Arrivé 1985). Nous dirons dans notre terminologie, qu'il n'est pas l'autre face du signifié, mais qu'il introduit un contenu pour quelqu'un.³⁵ Le signifié est lié à la représentation systémique de la langue alors que le contenu est lié à un acte. Certes il faut bien que les actes soient plus ou moins compatibles avec les représentations pour que cela fonctionne. Mais compatibilité ne signifie pas identité. . .

En un mot, un signifiant est donc un symbole devenu transparent par le fait qu'il introduit le sujet qui le discrimine inconsciemment à un nouveau contenu. Il s'ouvre sur une icône dans une nouvelle sémiose. Aussi le trait signifiant, par son détachement même, extériorise une expérience à laquelle il donne justement accès comme entité séparée. Autrement dit, si c'est l'aspect formel qui donne au signifiant sa saillance, cette forme enrichit l'univers du sujet, par la possibilité que s'y déposent les contenus des activités vécues en situation qui lui ont donné naissance. Un signifiant, à ce titre, représente le sujet dans sa quête de sens, au sens lacanien, car cette quête ne peut s'actualiser que par la reconnaissance de nouveaux signifiants. Ils ne se dévoilent cependant comme nouveaux signifiants que dans la mesure où ils arrivent au bon moment, déjà préparés par un usage ancien.

Reformulons: Tout signifiant, en tant que supporté, à l'origine, par un symbole, peut être interprété comme la marque de ruptures temporelles. La première marque de ce signifiant est pour Lacan le trait unaire. Par cette première marque formelle d'un acte, quelque chose de l'ordre temporel devient atemporel, devient structural. L'habitude active d'un retour et d'un certain contrôle sur ce retour en situation conduit à le reconnaître comme une même forme. Cela dit, la complexification des rythmes débouche progressivement à la conception de formes de plus en plus complexes.

En passant du symbole au signifiant, on passe progressivement d'une organisation temporelle de la complexité symbolique à une saisie de l'atemporalité, de la spatialité du signifiant propre à supporter des contenus. Le tissage, comme activité primitive, est peut-être une expression de ce passage allant d'une activité temporelle rythmée à la représentation spatiale affectée d'un contenu mythique.

Ce processus est également à la base de l'ordre du discours. Les symboles sont d'abord perçus par la distribution de leurs occurrences dans un temps, et c'est leur stabilisation complexe qui les introduit dans un nouveau médium qui n'est plus le temps, mais l'espace. On passe alors de l'acte et de son contenu à leur représentation par des signifiants. Les nombres entiers portent également la trace de ce passage de l'ordinal au cardinal.

L'ordre temporel engendre alors un ordre spatial. Si les signifiants émergent d'abord par leur détachement dans la conscience comme symboles en recouvrant des ruptures temporelles, les activités finissent par les associer à leurs produits, et notamment aux objets fabriqués. Le rythme de ces activités règle ainsi leur articulation qui visibilise de nouveaux produits, porteurs éventuels de nouveaux signifiants. . . porteurs également d'une logique de l'articulation, car les contenus et les actes qui, en définitive, les supportent, chez des sujets actifs, ne doivent pas être en contradiction trop flagrante avec le réel de ces productions.

Réciproquement — et c'est un aspect important au niveau de la méthodologie - comme pour le tissage, c'est le rythme des activités qui assure une lisibilité des signifiants. Le rythme permet le détachement des signes dans le temps d'une lecture. Selon ce point de vue, chaque signifiant nouveau implique un lien entre un contenu et des rythmes actifs, chez un sujet: ce dont il peut juger intuitivement. Cependant, l'ordre de leur lecture, de leur reconnaissance, implique une contrainte que le contenu à la fois épouse et rend lisible, d'où cette ambiguïté de la représentation, à la fois forme calculable et support d'une nouvelle icône.

Ainsi, l'organisation des signifiants repose sur un ordre qui peut finir par apparaître dans sa logique et sa calculabilité, mais on ne doit pas oublier qu'il est d'abord l'expression, par sa temporalité, des rythmes des activités d'un sujet.

La méthode ALCESTE

La modélisation du corpus

Nous proposons un modèle dans lequel est calculé un aspect de la signification non pris en compte par le sémanticien. Le discours y est conçu non pas par ce qu'on s'en représente mais par ce qui s'y inscrit comme activité.

On s'intéresse donc au discours dans un de ses aspects les plus archaïques avec l'hypothèse suivante: La signification d'un discours ne se forme pas essentiellement dans ce qui s'élabore à partir d'une représentation, puisque celle-ci ne peut être préalable à celui-ci; Elle s'élabore à partir d'une activité rythmique plus primitive, qui laisse une place centrale à la

répétition. C'est par la *répétition* qu'une stabilisation de l'activité discursive peut se concevoir. Mais cette *répétition* ne se dévoile dans sa complexité que dans un jeu ternaire. L'objectif de la méthode proposée est justement de modéliser un aspect du processus de nouage entre les trois brins iconique, indicelle et symbolique de la *répétition* afin de permettre une représentation statistique de la trame.

De même qu'une cartographie des reliefs ouvre à la possibilité de se représenter les rivières, les voies de communication, les activités sociales plus complexes, de même l'étude des *répétitions* dans un discours ouvre à la possibilité de se représenter les courants idéologiques, les zones de conflits, de ruptures, à partir de recouvrements, d'enveloppements ou d'oppositions entre ce que nous appelons des *mondes lexicaux* (1993).

Reprenons rapidement nos trois modes de *répétition* pour l'analyse statistique. On considère un discours dans sa temporalité comme une succession de moments discursifs sans donner pour l'instant à cette notion un statut bien précis. Formellement, on découpe le texte en segments disjoints successifs.

Chaque segment est supposé couvrir une faible durée. Il est donc la trace possible d'une *résonance iconique* (ou *contenu*). C'est en effet le propre d'une icône que de se prolonger par captation d'un regard ou d'un esprit dans une certaine durée.

La rupture entre deux segments de texte modélise, quant à elle, la faille de toute énonciation à installer durablement cette *résonance*. *Notons qu'elle ne peut modéliser une altérité réelle*, mais elle modélise la simple *possibilité* d'une telle altérité, dont on ne peut rien dire a priori, sinon que cela peut changer le cours du discours. De plus, le découpage du texte en segments permet de passer de la notion de *possibilité* à la notion de *probabilité*. Ce découpage ouvre en effet aux calculs, car la probabilité d'associer deux mots ne peut se calculer qu'une fois le texte découpé en segments disjoints (à partir de leurs cooccurrences dans les *mêmes segments*).

Formellement, ce découpage du texte, par l'alternance des moments pleins et des possibles instants de rupture, modélise un rythme de lecture. En effet, un acte de lecteur présuppose un rythme, et notre hypothèse est qu'un texte ne peut s'interpréter sans une scansion particulière, propre au lecteur ou au locuteur, en tant qu'il est *sujet*.³⁶ Si le modèle ne peut rien dire de cette scansion particulière, a priori, il peut cependant aider à la reconnaissance de nouveau *signifiant* dans la mesure où les résultats sont assez stables (pour le regard de l'analyste) relativement au type de découpage choisi. C'est ce que l'analyse statistique se propose justement de montrer par le calcul.

Le principe de la méthode est simple: le corpus à analyser est découpé en une suite de segments de texte et l'on observe la distribution des *mots pleins* dans ces segments. . . d'où le nom de la méthode: '*Analyse des Lexèmes Cooccurrents dans un Ensemble de Segments de Texte*'.

Comme on l'a déjà évoqué, ces deux types d'*unités* - lexèmes et segments - sont problématiques. Un *mot plein* n'est pas un *lexème*. De même un segment de texte n'est pas un énoncé. Nous avons vu que les notions de mots pleins et de coupures sont relatives à l'expérience d'un *sujet*, en fonction de sa propre perception des icônes (*contenus*). Aussi *il y a un aspect fondamentalement flou dans cette problématique de l'unité textuelle* aussitôt que l'on veut la saisir en dehors d'une expérience subjective. Cela n'implique cependant pas une impossibilité de calculer. Par exemple, la possibilité de représenter un objet concret n'implique pas d'avoir des *unités* précises. Prenons l'exemple d'une plaque photographique ou d'une photo numérique. L'image se forme à partir d'un grain. Le grain de la photo est défini de manière encore plus arbitraire que les *coupures* que nous proposons entre énoncés. Et cela importe peu dans la mesure où ce grain permet à un *sujet* de retrouver une *stabilité* dans les *formes* de son expérience (*signifiant*). La définition du grain est arbitraire et pourtant la possibilité de *représenter quelque chose* dépend de lui. Dans notre modèle, les grains sont les *segments de texte* ou *unités de contexte*³⁷ et la coloration des grains dépend des *mondes lexicaux*.

Une fois les *unités* approximativement retenues, on cherche à étudier la structure des cooccurrences des *mots pleins* dans les différentes *unités de contextes* (Cf. Tab. 1).

Ce tableau est une modélisation possible des trois aspects de la *répétition* associés à l'*objet* d'un discours pour un lecteur: 1) par ce qui est en lignes; 2) par ce qui est en colonnes; et 3) par ce qui se construit comme ordre à travers le tableau numérique proprement dit:

- a) *Répétition iconique*. La *répétition iconique* est modélisée par ce qui lie les colonnes entre elles. En effet, la présence simultanée de *mots pleins* dans une même *unité de contexte* est la trace possible d'un même acte de *contenu*. La coprésence de 'uns' dans une même ligne est donc une marque formelle probable de ce contenu.

Tab. 1. *Le tableau modélisant l'activité du locuteur*

Tableau de données	bonheur	corbeau	menu	mimuit	prison	rapace
<i>u.c. 1</i>	0	1	1	1	0	1
<i>u.c. 2</i>	1	0	0	0	1	0
<i>u.c. 3</i>	1	0	1	0	0	0

- b) *Répétition indicielle*. La *répétition* indicielle est modélisée par ce qui lie les lignes entre elles. Le recouvrement complet du texte par des segments de texte disjoints, les '*unités de contexte*', permet d'appréhender les différents 'moments' de l'énonciation. L'indicialité est donc modélisée ici, abstraitement, par sa mise en limite: la *coupure*³⁸ en tant que fin possible d'une répétition iconique (fin dont on sait qu'elle est l'expression d'une indicialité).
- c) *Répétition symbolique*. Les *coupures* entre les segments de textes, en tant qu'elles peuvent permettre de discriminer des *mondes lexicaux*, rendent compte d'un mouvement caractéristique des énoncés qui prennent alors sens dans une nouvelle configuration signifiante. Il se peut en effet que le système des *mondes lexicaux dégagé par le calcul tombe juste pour l'analyste*, autrement dit, *que ce système exprime ce que l'analyste est capable de concevoir du fait de sa propre expérience du monde*. En ce cas, il y a une forme de *résonance* entre ce qui est calculé et un usage vécu, entre une forme symbolique et une expérience particulière. C'est par cette résonance, que le calcul symbolique devient *signifiant, support d'une nouvelle icône* pour l'analyste. Cette résonance est simplement vécue par l'analyste comme *contenu* de l'analyse. La forme calculée devient alors transparente pour lui.

La Classification Descendante Hiérarchique (C.D.H.)

Le mode de calcul appliqué au tableau de données pour faire apparaître le système des répétitions s'inscrit, au plan technique, dans ce qu'on appelle 'l'analyse des données à la française' et dont le principal promoteur est J. P. Benzécri. L'algorithme que j'ai proposé dans ma thèse en 1979 et amélioré en 1986 (Reinert 1983, 1986) est adapté de techniques connues pour traiter un problème précis: celui de l'analyse de tableaux très vides (jusqu'à 99 % de zéros). Cette méthode est dérivée de l'analyse factorielle des correspondances. Elle permet aujourd'hui (version 5) d'analyser des tableaux ayant au maximum 40 000 lignes et 3 000 colonnes avec un nombre de 'uns', toutefois, qui ne dépasse pas 1 500 000. Ceci permet de traiter des corpus allant jusqu'à 60 millions de caractères.

Principe de la méthode. L'objectif est d'approcher les *mondes lexicaux* d'un corpus sans avoir à poser préalablement le problème de la définition de catégories de contenu. En effet, comme on l'a déjà évoqué, c'est la dynamique même du discours qui conduit à distinguer des ancrages topiques différents. *C'est donc parce que les unités de contexte ne s'ancrent pas dans les mêmes mondes lexicaux que ceux-ci vont finir par émerger*. Cet aspect différentiel de l'analyse statistique est, pour nous, fondamental au

niveau sémiotique, car *c'est lui qui permet de saisir dans la coupure même des classes la stabilisation d'un écart. C'est par cet écart que s'introduit l'idée seconde d'une proximité pour les éléments d'une même classe. Autrement dit, la représentation d'un contenu ne peut être que troisième (en aucun cas elle est première)*. Elle nécessite en effet d'avoir d'abord expérimenté la notion de *Différence*. Ainsi les fluctuations répétées de ce qui diffère permettent l'accès au niveau symbolique, ce niveau n'exprimant rien d'autre qu'une *logique des ruptures* dont les classes donnent justement une première représentation. C'est en cela que l'aspect différentiel de l'approche est essentiel pour comprendre le processus même de l'analyse.

À ce propos, il y a également *oscillation* entre ce qui se stabilise au niveau statistique et ce qui se normalise au niveau de ce qui est interprété. Les frontières entre classes statistiques vont peu à peu être appréhendées par l'interprète à travers une logique catégorielle. *Interpréter, c'est trouver un ajustement qui ne prouve rien en lui-même sinon qu'il permet d'apprécier concrètement ce qui tombe juste* entre ce que dit l'interprète et ce que montre l'analyse. *Cette appréciation, comme on l'a dit, n'est pas de l'ordre d'une preuve mais de l'ordre d'une abduction, d'une hypothèse, d'une résonance iconique.*

Principe du calcul. Le principe de la méthode de classification utilisée peut être présenté très simplement. L'objectif est de classer les segments de texte en fonction d'une qualité interne à ces segments: la liste des mots pleins qui y sont présents. Ce classement est indépendant de l'ordre entre segments ou de l'ordre entre mots si bien que l'on peut changer l'ordre des colonnes ou des lignes du tableau de données sans que cela modifie la structure que l'on cherche à mettre en évidence.

Cela étant, parmi tous les manières possibles d'ordonner les lignes et les colonnes, en existe-t-il une, qui permet de se rapprocher le plus de la disposition présentée Tab. 2, les deux sous-tableaux de la diagonale (en grisé) comprenant davantage de 'uns' que les sous-tableaux de l'antidiagonale (pratiquement vide).

Au plan du critère à maximiser, cette recherche de partitionnement passe par le calcul des *marges* de chacun des deux sous-tableaux considérés, étendu à l'ensemble des colonnes. Par exemple, la marge du tableau 1

Tab. 2. *La distribution recherchée 'idéale' pour l'obtention de la première dichotomie*

	Vocabulaire 1	Vocabulaire 2
Classe 1 de segments de texte	Tableau 1	
Classe 2 de segments de texte		Tableau 2

comptabilise, pour un mot du vocabulaire retenu (une colonne du tableau), le nombre de fois où ce mot est présent dans un des segments de texte de la première classe. Idem pour la marge du tableau 2 relativement à la classe 2. L'écart entre ces deux marges peut être mesuré à l'aide d'un χ^2 . Et le problème posé peut être maintenant rapidement formulé: *Quelle est la partition des segments de texte en deux classes qui maximise le χ^2 des marges?*

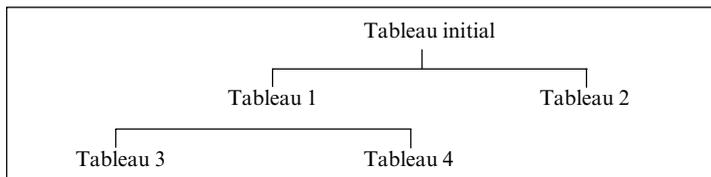
C'est-à-dire, quelle est la partition qui maximise la différence de profil entre les deux marges? Une fois cette partition obtenue, on peut déplacer à la gauche du tableau les mots relativement plus présents dans la classe 1 et à la droite du tableau les mots les plus présents dans la classe 2. D'où la représentation proposée Tab. 2.

Après ce premier calcul, il est possible de recommencer le même algorithme sur le plus grand des sous-tableaux restants à traiter, après élimination des colonnes presque vides ou liée à des mots très spécifiques de l'autre classe, etc. (cf. Tab. 3). D'où le nom de la méthode: la *Classification Descendante Hiérarchique*³⁹. Notons que cette méthode est, au plan des algorithmes, dérivée de l'analyse factorielle des correspondances (Benzécri 1973; Reinert 1983, 1986).

Conclusion

En réponse à la question initiale sur le rôle de l'approche statistique dans l'analyse du sens d'un discours, nous croyons avoir montré que cette approche est nécessaire par le fait même que le discours est l'expression d'une *répétition* en acte . . . sans pour cela qu'une telle analyse implique une maîtrise totale du sens, l'interprétation restant en définitive une affaire de sujet. Cela dit, la trace textuelle d'un discours est bien le chemin qui mène un locuteur ou un interprète à prendre conscience de quelque chose. Mais la linéarité de ce chemin est trompeuse car l'activité discursive est foncièrement *discontinue* et *fractale*. Seule la prise de conscience qui peut accompagner le franchissement d'une discontinuité conduit à concevoir cette dernière comme lien. La coupure devient alors symbole et renouvelle, par ce fait, le mouvement même de la sémiose à une autre échelle. Aussi le

Tab. 3. *L'arbre des analyses d'une Classification Descendante Hiérarchique*



contour de la sémiose ne peut être appréhendé d'un point de vue extérieur au sujet qui la vit. Elle dépend de l'état de conscience de ce sujet, de sa propre posture face à ce qu'il perçoit comme mode d'être.⁴⁰

Le problème de la dualité entre postures et modes d'être

De nombreuses discussions amicales avec P. Achard et P. Wald⁴¹ nous ont aidé à prendre conscience du rôle des postures, en analyse de discours. Dans une étude sur des témoignages d'anciens appelés de la guerre d'Algérie, P. Achard (1991) avait distingué trois *postures* dans les réponses des interviewés: la *posture de témoin*, la *posture d'acteur*, et la *posture de patient*.⁴² Ces *postures* avaient simplement été discriminées en fonction de la position de l'énonciateur dans la réponse, l'analyse étant purement discursive.

Deux ans après cette publication (que nous ne connaissions pas), nous proposons un article dans cette même revue sur l'analyse d'un corpus de récits de cauchemars (Reinert 1993). L'application d'*ALCESTE* à ce corpus se résumait par trois classes de récits et P. Achard fut surpris et intéressé d'y retrouver les traces des *postures* qu'il avait discriminées dans ses propres récits.

Aujourd'hui, nous proposons de voir dans ces *postures* l'expression subjective des *modes d'être* peirciens. Il est en effet assez naturel de penser que la *posture de témoin* est celle par laquelle l'attention du sujet est focalisée sur le *contenu immédiat de la situation* (priméité); Alors que, dans la *posture d'acteur*, le sujet privilégie son engagement dans la situation (secondéité). Quant à la *posture de patient*, elle met l'accent sur l'enchaînement subi des faits, . . . ce qui est une manière de se désengager en se représentant une loi (tiercéité).⁴³

Cela est dit rapidement, mais ce nouage entre *modes d'être* et *postures* rend compte — c'est en tout cas l'hypothèse de travail que nous soumettons — du mouvement même de la sémiose, *comme mode d'être impossible*; Impossible, car *il ne peut y avoir aucune extériorité à l'être pour en juger*. Mais, comme la dualité *sujet-objet* à laquelle elle s'apparente, cette dualité *posture-mode d'être* boîtie. Un déséquilibre l'ouvre au mouvement même de la sémiose, déséquilibre qui marque la place de ce que l'on ne peut pas maîtriser et qu'on appelle le temps.

La répétition et l'impossible maîtrise du signifiant

En conséquence, une '*sémiose*' est indéfinie puisqu'il ne peut y avoir de point de vue externe pour la saisir dans son *unité*. Il est impossible de dire,

par exemple, si un nouvel énoncé est vraiment un nouvel énoncé, s'il n'est pas simplement la suite d'un énoncé plus ancien; ou sa reprise, etc. *Ceci implique un lien de fondement entre sémiose et répétition*, qui prend appui sur la notion de *signifiant*, le signifiant étant la partie immergée de la sémiose, celle qui fait *écho* à une multitude d'autres moments. La prise de conscience caractérise justement le changement de sémiose par la reconnaissance d'un nouveau signifiant. Celui-ci surgit de l'obscurité des sémioses antérieures, mais il n'apparaît à la lumière qu'englué dans toute une matérialité à la fois corporelle, situationnelle et même culturelle que nous appelons justement *contenu*. Nous pourrions dire avec Lacan: '*Je parle avec mon corps et ceci sans le savoir. Je dis donc toujours plus que je n'en sais. C'est là que j'arrive au sens du mot sujet dans le discours analytique. Ce qui parle sans le savoir me fait je, sujet du verbe. Ça ne suffit pas à me faire être.*' (Lacan 1975: 150)

Un autre aspect posé par notre démarche est le suivant: En cherchant une représentation des *postures*, ne sommes-nous pas en train de faire ce que nous critiquons, à savoir vouloir définir une position objective absolue pour définir la sémiose? N'y-a-t-il pas là une contradiction? Nous ne le pensons pas car notre point de vue ne peut être métalinguistique. Il n'y a en effet aucune extériorité possible à la sémiose, et, en cela, nous sommes en accord avec Lacan lorsqu'il énonce qu'il n'y a pas de métalangage.⁴⁴ Si la recherche d'une maîtrise est un moment nécessaire de la sémiose, elle n'en est qu'un moment et traduit donc une illusion. Reconnaître l'existence de postures conduit seulement à l'acceptation de cette illusion comme exprimant la condition de *sujet*. Dans la danse, par exemple, le pas échappe à la volonté de maîtrise du danseur tout en y étant également soumis. Il est cependant autre par rapport à son attente. A chaque fois il surprend. C'est justement par le jeu changeant des *postures* que cette contradiction de l'acteur s'assume dynamiquement, en se complexifiant, en s'approfondissant, en lui donnant des résonances nouvelles et inconnues. Le danseur n'arrive pas à danser comme il voudrait mais il danse avec son style, mêlant acte, représentation et plaisir de la mimesis. Au niveau d'un discours, c'est par ce jeu tournant des postures qu'un impossible à dire finit par se montrer *autrement*. Ce quelque chose de la répétition advient alors comme une pure création. Comme Deleuze le commente à propos de la philosophie de Nietzsche, l'acte de création *affirme* la *répétition* dans sa nouveauté irréductible. . . Création qui prend cependant sens dans la logique de *répétition* d'une œuvre ou d'un parcours singulier; *Logique de sujet*,⁴⁵ disons-nous, à partir de notre lecture de Lacan, qui s'affirme contre une *logique représentée*, normative, établie, qui s'affirme par la reconnaissance active de ce qui advient. Au plan du discours, tels sont les énoncés qui

s'échangent, comme des pas de danse, toujours nouveaux et toujours les mêmes. D'où cette citation de Nietzsche en exergue: '*C'est une douce folie que le langage: grâce à lui l'homme passe en dansant sur toutes les choses*'.

Abduction et logique du sujet

C'est parce que ça se répète qu'un sujet s'affirme comme existant. Pour Nathalie Charraud (1997): 'En psychanalyse, ce qui limite le hasard n'est pas la notion de probabilité, mais celle de répétition. Les événements qui arrivent comme par hasard sont pris dans un processus de répétition: quelque chose se répète, au moyen de la réalité'. D'un point de vue sémiotique et logique, ce quelque chose est justement saisi par abduction. Nous situons là cette possibilité de l'analyse statistique du discours à appréhender, à travers le jeu des répétitions, la trace d'un réel, dont le sujet est en quelque sorte le jouet et le médium; Se forme ainsi, à travers la configuration même de la trace de son activité, ce qu'il cherche justement à se représenter et à communiquer par son propre discours . . . Discours qui devance en quelque sorte sa propre conscience.

Au plan de la méthodologie, la limite formelle de ce type d'approche est justement de savoir s'arrêter à la porte de tout contenu, de tout imaginaire, et surtout de tout signifiant.⁴⁶ Car celui-ci ne peut-être appréhendé comme tel, c'est-à-dire comme forme séparable, que par un sujet, dans le mouvement même des pas d'un désir qui le fait tenir debout dans sa traversée des signifiants, et poursuivre son chemin.

Notes

1. Une toute première ébauche de cette étude a été présentée au colloque de Cerisy organisé par M. Arrivé et C. Normand en septembre 1998 et intitulé '*Linguistique & Psychanalyse*' (Reinert, 2001a). Ce travail a principalement été effectué dans le cadre du C.N.R.S.. Il a également bénéficié d'une subvention du Ministère de la Recherche dans le cadre d'une Action Concertée Incitative: ACI Cognitive-13b, dirigée par M.C. Noël-Jorand (1997).
2. Extrait de '*Ainsi parlait Zarathoustra*' traduit par G. Goldschmidt (Nietzsche 1972).
3. Un discours, en tant qu'il n'existe que pris en charge par un locuteur, ici et maintenant, ne peut avoir qu'un sens. . . Celui que ce locuteur poursuit. Ce n'est bien sûr pas le cas du texte qui le soutient.
4. Jean-Paul Benzécri est l'inventeur de l'*analyse factorielle des correspondances* comme méthode d'analyse de 'données linguistiques' qui date de 1962 (Benzécri 1982). Rappelons que cette méthode a été conçue à l'origine comme méthode 'inductive' d'analyse statistique des textes. Le problème sémiotique posé par cette approche a déjà été évoqué par certains auteurs (Bolasco 1982).

5. Le 'de' d'analyse 'de' discours, cher à P. Achard, signifie pour nous qu'il ne peut y avoir *analyse 'du' discours*, pour un analyste prenant comme support de son analyse un texte, car la textualité d'un discours l'ouvre à une multitude de discours possibles.
6. Nous avons hésité à reprendre le terme polémique de Pêcheux (1969) d' *'analyse automatique du discours'*. Mais le 'du' est trop problématique. Alceste simule cependant une analyse de discours car il modélise, non pas le texte, mais l'activité d'un lecteur par le découpage du texte en unités de contexte.
7. Nous appelons *locuteur* celui qui prend en charge le discours. . . Notons qu'un lecteur, un auditeur ou même un analyste, doit prendre également en charge le discours quand il s'engage dans un sens.
8. Titre donné à la publication de son *séminaire XI* (Lacan 1964).
9. Deuxième Lettre à Lady Welby (Peirce 1978, p.51).
10. Michel Balat (2000) a recherché chez Peirce des fondements sémiotiques à la psychanalyse.
11. C'est un des handicaps de la sémiologie issue de l'approche saussurienne que d'avoir voulu définir le signe à partir d'unités minimales formelles.
12. D.-R. Dufour (1988: 30) évoque une proximité de ce type entre cette définition de Lacan et celle de Lévi-Strauss, appliquée au récit, et qu'il formule ainsi: *'Une version [d'un récit mythique] est ce qui représente le récit pour une autre version'*.
13. [CP 2.228, 1897]. Les références aux *Collected Papers (CP)* sont de G. Deledalle.
14. À ne pas confondre (bien qu'il y ait un rapport) avec la *logique naturelle* de J. B. Grize (1990) qu'il appelle également *logique de sujets* (avec un 's') car dialogique et sociale (p. 21). La *logique de sujet* dépend de la manière dont l'histoire particulière d'un sujet est vécue en situation. Elle découle plus directement de la conception du *sujet* chez Lacan comme *effet du signifiant*: *'Le sujet, ce n'est rien d'autre - qu'il ait ou non conscience de quel signifiant il est l'effet - que ce qui glisse dans une chaîne de signifiants.'* (1975, p. 65). Voir aussi note 46.
15. Pour Peirce, l'objet dynamique est l'objet en dehors du signe par opposition à l'objet immédiat présenté dans le signe. Cet objet dynamique, *'le signe doit le présenter par suggestion; et cette suggestion ou sa substance, c'est l'objet Immédiat'* (p. 53). Notre version est plus lacanienne. Elle diffère par la place que nous donnons au *sujet* mais elle reste compatible avec cette interprétation si l'on accepte l'idée que l'aspect dynamique de l'objet, lié à un usage réel des signes, présuppose un *sujet* pour qui cela est bien ainsi.
16. CP 8.335.
17. CP 2.304.
18. Car la notion de *ressemblance* implique une comparaison, et donc un jugement (tiercéité).
19. Saillance que l'on peut concevoir comme résonance avec des engrammes? Dans l'homme neuronal (Fayard 1983), Changeux évoque ce problème de la *résonance* en relation avec les propriétés oscillatoires *'des assemblées de neurones'* (p. 183): *'La discussion théorique sur la genèse des objets mentaux et sur leur enchaînement a mis en relief une opération critique qui donne aux 'calculs' cérébraux leur sens: l'épreuve de la réalité. Il a été proposé que l'entrée en résonance (Thom 1980) d'objets mentaux, ou leur dissonance pouvaient en rendre compte. Peut-on imaginer un mécanisme cellulaire d'entrée en résonance des assemblées de neurones?'*. Par cette citation, on ne cherche pas à faire du réductionnisme, mais on désire seulement suggérer que ces différents modèles ne sont pas incompatibles entre eux.
20. On reviendra largement sur l'aspect *signifiant* d'une icône quand on envisagera la notion de *symbole*.

21. À approfondir relativement aux conceptions de l'Inconscient de Freud (lettre 52 à Fliess). Nous reprenons ce problème de la prise de conscience dans notre conclusion.
22. L. Zapata (2001) nous a fait remarquer l'importance de cette notion de style chez Lacan qui commence ses *Écrits* (1966) par cette phrase de Buffon: '*Le style c'est l'homme même*'.
23. Sur la notion d'isotopie, se référer à Greimas et Rastier (Rastier 1972).
24. CP 2.248.
25. L'indexicalité de l'énoncé est sans aucun doute en rapport avec l'indexicalité (au sens du *dictionnaire d'analyse du discours*: cf. Charaudeau et Maingueneau, 2002). Cela dit, ces deux termes ne se confondent pas. Un indice est en rapport avec ce moment où l'icône devient étrange, alors qu'un indexical est une marque que l'on pose sur la coupure, là où il y aurait un trou dans le discours. Par ce processus, l'origine de l'énonciation devient visible. Par ce fait, un indexical est déjà un symbole au sens de notre &1.3.
26. Voir par exemple l'article 'polyphonie' du dictionnaire d'analyse du discours (O.c.)
27. Voir l'article 'discours' de l'Encyclopaedia Universalis (1996, version 2.0.)
28. Selon Mendelbrot, la notion de *fractale* reste elle-même ambiguë, dans sa naturalité. Au plan mathématique, on l'approche à l'aide de fonctions récursives (Mendelbrot 1975)
29. CP 2.292.
30. Pour G. Mounin et P. Encrevé, l'habitus, en sociolinguistique, se définit comme: '*Système de dispositions socialement structurées et structurantes, qui règle l'acquisition et l'utilisation du langage*' (article 'sociolinguistique' de l'E.U., o.c.). Elle pêche par la non prise en compte du *sujet* qui conduit à voir dans l'habitus, une forme réglée, c'est-à-dire, représentable. Nous utiliserons le terme d'habitude en tant qu'il désigne un vécu répétitif immédiat; Un individu ne peut en avoir une conscience absolue, même si la conscience qu'il en a, dépendant justement des habitus sociaux. Il est clair que si les habitus déterminent des habitudes, les habitudes influent également sur les habitus. L'habitus est d'ailleurs plus qu'un simple système de règles. Chaque individu ne peut en avoir qu'une représentation simplifiée, partielle, mythique. Et surtout, personne ne peut en avoir une représentation absolue. L'acteur en sait toujours plus qu'il ne se représente.
31. La différence type-token a été posée par Peirce.
32. O.c., article 'symbole'.
33. Sur cette notion de tombé-juste, j'ai été directement influencé par les conceptions du mathématicien et philosophe, René Guitart (1999) dont j'ai suivi un séminaire dans le cadre d'un groupe de travail de la revue *Langage Société* (Maison des Sciences de L'Homme, Paris).
34. Par opposition au prédicat, qui dit quelque chose sur ce qui est visé par le sujet.
35. L'icône est perçue par mimesis (résonance) et non pas par représentation. Elle est perçue comme *contenu* d'un acte. La différence peut sembler difficile à apprécier. Elle ne l'est pas, en pratique, car la mimesis est une expression immédiate d'un contenu alors que la représentation est une expression médiata qui prend appui sur un assemblage de signifiants distincts: mots, règles, conventions, lois, calculs, scénario, concept, schéma. Cela dit, ce qui est de l'ordre de la représentation à un moment peut, par apprentissage ou par habitude, apparaître à un autre moment, comme une icône, quand le sujet ne distingue plus les différents signifiants. L'ambiguïté de la notion de représentation est une des raisons de son abandon par Peirce pour définir la tiercéité. On voit également cette ambiguïté chez Freud quand il distingue, 'représentation d'objet' de 'représentation de mot'.
36. C'est en cela qu'ALCESTE propose une modélisation de l'activité du locuteur-lecteur et non une modélisation du texte, en tant qu'il signifierait de lui-même.

37. Cette propriété des analyses 'ALCESTE' à rester relativement stables quel que soit le 'grain' choisi est un autre indice pour défendre l'hypothèse d'une fractale de l'énonciation. On sait qu'une fractale conserve une forme stable quelle que soit l'unité prise pour la décrire. Dans son livre sur 'les objets fractals', Mandelbrot (1975) donne l'exemple de la côte de Bretagne qui garde une même structure qu'elle soit représentée au 1/100000 ou au 1/1000. L'aspect fractal tient dans le fait que toute partie est homothétique au tout (page 25). C'est d'ailleurs cette propriété qui permet de construire des formes fractales à partir d'une fonction récursive.
38. Cela dit, nos coupures tiennent compte de la ponctuation, si elle existe, afin d'intégrer ces coupures à une scansion *possible* pour un lecteur quelconque, même si elles restent fondamentalement arbitraires.
39. Cette terminologie est utilisée par d'autres auteurs et recouvre parfois des techniques différentes (voir Benzécri 1973: 1, 315–319). La technique que je présente (Reinert 1983, 1986) est celle que j'ai mise au point et présentée dans ma thèse avec Benzécri pour directeur en 1979 (Université Paris VI).
40. Les modalités de l'être ont fait couler beaucoup d'encre depuis Aristote. Peirce, quant à lui, distingue trois modes d'être phénoménologiques qu'il appelle: priméité, secondéité et tiercéité. Ils se traduisent au niveau de la saisie d'un signe, dans son rapport à l'objet, par la différenciation: *icône, indice et symbole*.
41. Dans le cadre du séminaire 'analyse de discours' organisé par la revue *Langage Société*, Maison des Sciences de l'Homme, Paris. Cf. Achard 1993, Wald 1999.
42. Achard a mis en évidence 'trois postures de base suivant le procès posé au centre' de la réponse et la relation de 'je' à ce procès: — le procès de base n'implique pas 'JE' (position de témoin); — comme notion source ('sujet profond') du procès (position d'acteur); — comme notion but ('objet profond' ou autre place régie) du procès (position de patient)' (*Langage Société*: n°55, p22). Il cite un exemple de réponse qui reprend les trois postures (que nous mentionnons entre crochets): 'J'hésite entre: 1/ la mort d'un camarade [posture de témoin?]; 2/ le souvenir d'une lâcheté commise pour ne pas aller en opérations [posture d'acteur]; 3/ une semaine de Sirocco et de dérèglements de tous ordres [posture de patient]'.
43. M. Balat a consacré le chapitre 2 de son livre (O.c., p. 37–67) au lien entre les registres de Lacan et les catégories phénoménologiques de Peirce, avec l'équivalence suivante: priméité-imaginaire, secondéité-réel, et tiercéité-symbolique. Ce qui nous semble assez évident et proche même du vocabulaire de Peirce lui-même. Nous pourrions dire, que les registres de Lacan reprennent les catégories de Peirce sur le mode Imaginaire; alors que les catégories de Achard, les reprennent sur le mode du Réel. Les catégories de Peirce, étant présentées par lui, sur le mode purement symbolique . . . On ne peut parler des catégories que par métaphores car il n'y a pas de métalangage, pas de posture universelle pour cela.
44. C'est une proposition que Lacan a toujours maintenue: 'Je soutiens toujours qu'il n'y a pas de métalangage. Tout ce qu'on peut croire être de l'ordre d'une recherche du méta dans le langage, c'est simplement, toujours, une question sur la lecture' (Sem. XVII, 219).
45. Nous préférons toutefois 'logique de sujet' à 'logique du signifiant' car cette dernière semble présupposer une instance déterminée en dehors du sujet alors que le signifiant 'n'existe' que par le mouvement même d'un sujet qui glisse d'un signifiant à l'autre. N'oublions pas que dans notre conception, qui est, nous semble-t-il, celle de Lacan, le signifiant est simplement porté par le temps d'une coupure, en tant qu'elle est vécue comme telle par un sujet, en situation (en tant qu'elle fait résonance).

46. D'où la nécessité d'un découpage relativement arbitraire du texte en segments car rien ne permet de justifier qu'un segment de texte, défini de l'extérieur, représente bien un *signifiant* pour quelqu'un.

Références

- Achard, Pierre (1991). Une approche discursive des questionnaires: l'exemple d'une enquête pendant la guerre d'Algérie. *Langage et Société* 55, 4–40.
- (1993). *La Sociologie du Langage*. Paris: P.U.F.
- (1997). L'engagement de l'analyste à l'épreuve d'un événement. *Langage et Société* 79, 5–38.
- Arrivé, Michel (1985). Signifiant saussurien et signifiant lacanien. *Langages* 77, 105–116.
- Authier-Revuz, Jacqueline (1982). Hétérogénéité montrée et hétérogénéité constitutive: élément pour l'approche de l'autre dans les discours. *DRLAV* 26, 91–151.
- Bachelard, Gaston (1942). *L'eau et les rêves*. Paris: Librairie José Corti.
- (1949). *La psychanalyse du feu*. Paris: Gallimard.
- Bakhtine, Michael (v.f. 1977). *Le marxisme et la philosophie du langage*. Paris: Éditions de Minuit.
- Balat, Michel (2000). *Des fondements sémiotiques de la psychanalyse*. Paris: L'harmattan.
- Benveniste, Emile (1974). *Problèmes de linguistique générale II*. Paris: Gallimard.
- Benzécri, Jean-Paul et col (1973). *Analyse des Données*, Tomes 1 et 2. Paris: Dunod.
- (1981). *Pratique de l'Analyse des Données: linguistique et lexicologie*. Paris: Dunod.
- (1982). *Histoire et préhistoire de l'analyse des données*. Paris: Dunod.
- Bolasco, Sergio (1982). Vers une interprétation sémiotique de l'analyse des données. *Metron* vol XL, n° 1–2, 93–102.
- Charaudeau, Pierre et Maingueneau, Dominique (2002). *Dictionnaire d'analyse du discours*. Paris: Seuil.
- Charraud, Nathalie (1997). *Lacan et les Mathématiques*, Paris: Economica (Anthropos).
- Deleuze, Gilles (1970). *Nietzsche et la philosophie*, Paris: P.U.F.
- Dufour, Dany-Robert (1988). *Le bégaiement des Maîtres, Lacan, Benveniste, Lévi-Stauss...* . . . Strasbourg: Arcanes.
- Eco, Umberto (1976, v.f. 1992, Bouzaher). *La production des signes*, Paris: Le Livre de Poche.
- Foucault, Michel (1971). *L'ordre du discours*. Paris: Gallimard.
- Freud, Sigmund (1950, v.f. 1956, A. Berman). *La naissance de la psychanalyse*. Paris: P.U.F.
- Grize, Jean-Blaise (1990). *Logique et langage*. Ophrys.
- Guitart, René (1999). *La pulsation mathématique*. Paris: L'Harmattan.
- Harris, Zellig S., (1954, v.f. 1969, Dubois-Chalier). 'Analyse du discours', *Langage*, N° 13, 8–45.
- Jenny, Jacques (1997). Méthodes et pratiques formalisées d'analyse de contenu et de discours dans la recherche sociologique française contemporaine. Etat des lieux et essai de classification. *Bull. de Méthodologie sociologique* 54, 64–122.
- Lacan, Jacques (1966). *Ecrits*. Paris: Seuil.
- (1964, v.p. 1973, J.-A. Miller). *Le Séminaire, Livre XI (Les quatre concepts de la psychanalyse)*. Paris: Seuil (Points).
- (1969–1970, v.p. 1991, J.-A. Miller). *Le Séminaire, Livre XVII (L'envers de la psychanalyse)*. Paris: Seuil.
- (1972–1973, v.p. 1975, J.-A. Miller). *Le Séminaire, Livre XX (Encore)*. Paris: Seuil (Points).
- Lalande, André (1926, 3ème ed. 1993). *Vocabulaire Technique et Critique de la Philosophie*. Paris: P.U.F.

- Lebart, Ludovic et Salem, André (1994). *Statistique Textuelle*. Paris: Dunod.
- Mandelbrot, Benoît (1975, 4ème ed. 1995). *Les Objets Fractals*. Paris: Flammarion.
- Mircéa, Eliade (1969). *Le Mythe de L'éternel Retour*. Paris: Gallimard.
- Noël-Jorand M. -C., Reinert M. et al. (1997). A new approach to discourse analysis in psychiatry, applied to a schizophrenic patient's speech. *Schizophrenia Research Elsevier Science* 25, 183–198.
- Nietzsche (v.f. 1972, traduit par G. Goldschmidt). *Ainsi parlait Zarathoustra*. Paris: Le livre de poche.
- Pêcheux, Michel (1969). *Analyse automatique du discours*. Paris: Dunod.
- Peirce, Charles S. (1931–1935). *Collected Papers of Charles Sanders Peirce*, 8 vols., Charles Hartshorne, Paul Weiss, and A. W. Burks (eds.). Cambridge, MA: Harvard University Press.
- (v.f. 1978, Traduit et commenté par G. Deledalle). *Écrits sur le signe*. Paris: Éditions du Seuil.
- Rastier, François (1972). Systématique des isotopies. In *Essais de sémiotique poétique*, Greimas (ed.), 80–106. Paris: Larousse.
- Reinert, Max (1983). Une méthode de classification descendante hiérarchique. *Cahiers de l'Analyse des Données* 3, 187–198.
- (1986). Un logiciel d'analyse lexicale: ALCESTE. *Les cahiers de l'Analyse des Données* 4, 471–484.
- (1990). ALCESTE, une méthodologie d'analyse des données textuelles et une application: Aurélia de Gérard de Nerval, *Bulletin de Méthodologie Sociologique* 26, 24–54.
- (1993). Les 'mondes lexicaux' et leur 'logique' à travers l'analyse statistique d'un corpus de récits de cauchemars. *Langage et Société* 66, 5–39.
- (2001a). Processus catégorique et co-construction des sujets et des mondes à travers l'analyse statistique de différents corpus. In *Linguistique et Psychanalyse*, [colloque de Cerisy, Sept. 1998], M. Arrivé et Cl. Normand (eds.), 379–392. Editions in Press.
- (2001b). ALCESTE, une méthode statistique et sémiotique d'analyse de discours — Application aux 'Rêveries du promeneur solitaire', *Revue Française de Psychiatrie et de Psychologie Médicale* n° 49, 32–36.
- Saussure, Ferdinand (1972). *Cours de linguistique générale*. France: Payot.
- Wald, Paul (1999). Classes d'énoncés, dimensions modales et catégories sociales dans ALCESTE, *Utinam*, 1999–1/2, 303–24.
- Wittgenstein Ludwig (1961, v.f. Klossowski). *Tractatus logico-philosophicus*. Paris: Gallimard
- Zapata, Luz. et Sauret, Marie-Jean, (2001). Analyse du discours et méthode psychanalytique: la question du style dans la clinique des névroses, *Revue Française de Psychiatrie et de Psychologie Médicale* 49, 57–65.

ANNEXE. Quelques résultats de l'analyse du texte de cet article par 'ALCESTE'.

On a ôté du texte la bibliographie, les références, les notes, ainsi que le &2.22 dédié à la présentation de l'algorithme. Les tableaux de résultats présentés sont issus de l'analyse standard d'Alceste avec la version 4.5 du logiciel.¹ Voici un extrait du texte retranscrit pour l'analyse (Cf. Tab. 1a).

Tab. 1a. *La retranscription du corpus*

<p>**** *Introduction</p> <p>L'objectif que nous nous proposons est d'aborder la notion de signe de manière très générale en partant de la notion de répétition, afin de suggérer que l'approche du sens des discours dans certains de ses aspects ne présuppose pas nécessairement l'utilisation de modèles d'analyse de la signification linguistique. (. . .)</p> <p>répétition masquée'. Il ajoute: 'Le nouveau n'est pas dans ce qui est dit, mais dans l'événement de son retour'.</p> <p>**** *P_1 *1.0</p> <p>1.01. La notion de répétition est sans doute trop fondamentale pour sortir du langage courant.</p>
--

Une analyse standard enchaîne quatre étapes de calcul dont nous présentons un bref résumé dans les 4 paragraphes qui suivent.

A. Le corpus est découpé en fragments de texte, les unités de contexte élémentaires (u.c.e.), et les mots sont comptabilisés (Cf. Tab. 2a).

A la fin de cette première étape, le tableau de données modélisant l'activité discursive est constitué. Il croise les 222 unités de contexte élémentaires (u.c.e.) avec les 217 mots pleins (Cf. Tab. 2a). A l'intersection d'une ligne et d'une colonne, la valeur '1' signifie que le mot est présent dans l'u.c.e., et la valeur '0', qu'il est absent.

B. Les deux Classifications Descendantes Hiérarchiques (C.D.H.).

A partir du tableau précédent, deux autres tableaux sont calculés en concaténant les u.c.e. successives (dans le texte) de façon à obtenir des unités de contexte (u.c.) plus longues, chaque u.c. devant comprendre un nombre minimum imposé de mots différents analysés: Ici, 10 pour le tableau

Tab. 2a. *Les caractéristiques générales du corpus*

Nombre total de formes distinctes (mots non lemmatisés)	1672
Nombre d'occurrences	8559
Fréquence minimum d'un 'mot' pour l'analyser	5
Nombre d'occurrences par u.c.e.	27
Nombre de mots pleins analysés	217
Nombre de mots outils retenus	109
Nombre d'unités de contexte élémentaires (u.c.e.)	222

Tab. 3a. *Les caractéristiques des deux tableaux soumis à la C.D.H.*

Premier tableau (Donn.1)	Nombre minimum de mots par u.c.: 10	Nombre d'u.c.: 140
Second tableau (Donn.2)	Nombre minimum de mots par u.c.: 12	Nombre d'u.c.: 121

DONN.10 et 12 pour le tableau DONN.12 (Cf. Tab. 3a). Chacun de ces deux tableaux est ensuite analysé par la C.D.H., puis les résultats sont comparés en prenant comme unité, l'u.c.e.. En effet, toute u.c. de DONN.10 ou de DONN.12 est composée d'un nombre entier d'u.c.e.. Par conséquent une classe d'u.c. de l'une ou de l'autre analyse peut toujours être considérée comme une classe d'u.c.e..

Les cases diagonales du tableau Tab. 4a. informent sur la stabilité d'une partition en trois classes. Cette partie stabilisée regroupe ici 85.59 % des u.c.e. du corpus (64+60+66=190 sur 222). On dégage ensuite le profil de ces classes stabilisées, les calculs étant limités aux seules 190 u.c.e. stables.

C. Profil des classes stabilisées

Les mots pleins (Cf. Tab. 5a) sont ordonnés en fonction d'un coefficient d'association (χ^2 à 1dl). La différence de tonalité entre les mondes lexicaux est sensible. Dans cette analyse, le mot le plus spécifique pourrait être sélectionné comme mot-pivot: 'objet' pour la classe 1; 'signifiant' pour la 3; Et 'analyse' pour la 2. Cette classe 2 est la plus caractéristique: Elle se distingue dès le premier pas de la C.D.H. dans les deux analyses (d'où l'arbre présenté Tab. 5a). Son vocabulaire spécifique suggère des préoccupations plus pragmatiques en relation avec la méthodologie d'analyse statistique des textes. Quant aux deux autres classes, l'une (Classe 1) évoque un ancrage sémiotique traditionnel (*objet, signe référent*). Et l'autre (Classe 3) se distingue par des références à une logique temporelle: *ordre, temporel, action, logique, devenir, lecture, rythme, acte*, etc. Le &1.3 relatif à la *répétition symbolique* est nettement discriminée par cette classe (Cf. Tab. 7a). Autrement dit, ce paragraphe se distingue nettement des autres paragraphes de la partie 1. Cela souligne la spécificité de notre

Tab. 4a. *Tableau croisant les deux partitions obtenues (en nombre d'u.c.e.)*

Partition 1 (C.D.H. de Donn.1) Partition 2 (C.D.H. de Donn.2)	Classe 1	Classe 2	Classe 3
Classe 1	64	7	18
Classe 2	1	60	1
Classe 3	2	2	66

Tab. 5a. Liste des mots pleins les plus spécifiques des classes²

Classe 1 (64 u.c.e.)	χ^2	Nb uce	Classe 3 (66 u.c.e.)	χ^2	Nb uce	Classe 2 (60 u.c.e.)	χ^2	Nb uce
objet+	47	28	signifiant+	28	21	analyse+	36	17
signe+	41	27	nouveau+	21	11	texte+	25	11
lieu+	14	7	symbole+	15	16	statisti<	25	11
fondement+	12	6	lien+	11	6	methode+	20	9
matiere+	12	6	ordre+	11	9	unite+	18	8
peircien+	12	6	trait+	11	6	modelis+	18	8
tenir.	11	7	tempor+el	11	6	plein+	16	9
feu+	10	5	act+ion	9	14	discours	14	17
rapport+	10	5	log+ique	9	9	propos+er	14	8
exist+er	10	5	devenir.	7	8	permettre.	13	9
impos+er	10	5	distingue+	6	5	possi+ble	13	10
refer+ent	10	5	lecture+	6	5	significati+f	11	5
idee+	8	7	rythme+	6	5	aspect+	11	10
repet+er	8	7	passer	6	5	trac+er	11	8
sujet+	6	19	nouvel+	5	7	context<	11	5
definition+	6	8	acte+	5	10	hypothe<	11	5
appel+er	6	6	contenu+	5	17	corpus	11	5
parl+er	6	6	donn+er	5	7	monde+	9	7
sembl+er	6	6	conscienc+e	4	6	entre	9	11
naitre.	6	3	fait	4	11	simple+	7	7
indice+	5	7	representat+ion	3	8	experience+	7	6
notion+	3	9	dire+	2	6	enonce+	7	12
re+el	3	6	usage+	2	5	approch+e	5	5

Tab. 6a. Liste des mots outils les plus spécifiques des classes

Classe 1 (64 u.c.e.)	χ^2	Nb uce	Classe 3 (66 u.c.e.)	χ^2	Nb uce	Classe 2 (60 u.c.e.)	χ^2	Nb uce
pour	11	29	toujours	5	3	en-effet	7	7
son	9	18	ainsi	5	8	on	4	15
elle	7	16	alors	5	6	a-partir-d<	3	5
autre+	7	12	cependant	5	8	a-travers	3	5
non	6	6	sur	4	9	y	2	5
devoir.	5	6	qui	4	23	sont	2	5
mais	5	17	et	3	32	ces	2	6

Tab. 7a. Liste des parties et paragraphes spécifiques des classes

Classe 1 (64 u.c.e.)	χ^2	Nb uce	Classe 3 (66 u.c.e.)	χ^2	Nb uce	Classe 2 (60 u.c.e.)	χ^2	Nb uce
*1_0	65	31	*1_3	71	37	*P_2	85	34
*P_1	41	62	*P_1	5	51	*2_1	62	25
*1_1	12	16	*3_2	2	6	*2_2	16	9
			*Conclusion	2	11	*Introduction	3	5

élaboration sur la *répétition symbolique* qui prend aussi appui sur la notion de signifiant chez Lacan.

Les contrastes entre classes expriment ainsi la pluralité de nos ancrages théoriques: Influence peircienne pour la classe 1; Influence peut-être plus lacanienne pour la classe 3? Et statistique textuelle pour la classe 2.

Notons, pour terminer, que les différents *mondes lexicaux* mis en évidence par une analyse discriminent souvent des modes d'argumentation. Il y a comme une métaphore vivante du contenu par le style de l'argumentation. Même si, ici, il est difficile de l'évaluer (Cf. Tab. 6), il semble cependant que la classe 2, qui repose plus sur l'approche statistique, utilise un mode d'argumentation plus logique (*en effet, à partir de*). La classe 1 qui fait référence à la théorie peircienne utilise une argumentation plus dialogique (*pour, non, mais, devoir*). La Classe 3, enfin, avec son interrogation plus spécifique sur le temps, sélectionne des marqueurs logiques imprégnés de temporalité (*toujours, ainsi, alors, cependant*). C'est cet aspect redondant, répétitif, mimétique qui constitue le fondement topique d'une classe.

D. Présentation des 4 u.c.e. les plus spécifiques de chaque classe (Cf. Tab. 8a)

1. Version MacIntosh personnelle correspondant à la version 4.5 diffusée par la société IMAGE.
2. La colonne 'Nb uce' indique le nombre d'u.c.e. différentes de la classe qui contient le mot; Les valeurs du χ^2 sont à un degré de liberté. L'unité de calcul est l'u.c.e. Les signes '+' indiquent une lemmatisation possible (i.e. une réduction à la racine).
3. Le χ^2 est à 1 degré de liberté. L'unité statistique est l'occurrence d'un mot. Par exemple, l'u.c.e. n°35 a été sélectionnée car elle appartenait déjà à la première classe et qu'elle contient beaucoup d'occurrences de mots spécifiques de cette classe. Les mots spécifiques de la classe sont précédés d'un '#'.

Tab. 8a. Les 4 u.c.e. les plus spécifiques³ de chacune des trois classes

N°	χ2	Texte de l'u.c.e. sélectionnée dans la classe 1 ('objet')
35	36	il #tient #lieu de cet #objet, non sous tous #rapports, mais par #référence à une #sorte d' #idée que j'ai #appelée quelquefois le #fondement du representamen.
37	26	C'est en tout cas, notre #interprétation de'il #tient #lieu de cet #objet, par #référence à une #sorte d #idée', qui était déjà dans ce qui #fonde le #premier #signe, du fait des sémoses antérieures.
80	26	l' #indice, au sens #peircien, #évoque au-contre un #existant, hic et nunc, dont on ne #parle pas directement mais qui s' #impose justement à travers son #indice. plus précisément, pour Peirce, un #indice est un #signe qui #renvoie à l' #objet qu'il denote parce qu'il est #réellement affecté par cet #objet.
39	16	partant de cette #définition, Peirce distingue ensuite trois #modes de fonctionnement d' un #signe relativement à son #objet: l' icône, l' #indice et le symbole. Très succinctement, l' icône #évoque son #objet en vertu de sa nature #interne; l' #indice est #réellement déterminé par son #objet comme, par exemple, la fumée par rapport au #feu;
N°	χ2	Texte de l'u.c.e. sélectionnée dans la classe 3 ('signifiant')
146	29	le #rythme de ces #activités #règle ainsi leur articulation qui visibilise de #nouveaux produits, porteurs éventuels de #nouveaux #signifiants, porteurs également d' une #logique de l articulation, car les #contenus et les #actes qui, en définitive,
105	17	ainsi cette coupure #devient #lien. C'est par elle qu' on va pouvoir accéder à une #nouvelle icône, qui est le #contenu de la #proposition, et qui, dans une #proposition #logique, se confond avec le #prédicat.
120	17	1.33 cette définition du #symbole rejoint-elle la définition du #signifiant chez lacan et plus précisément, du #trait #unaire, par lequel, un #acte #devient objet de #représentation?
135	17	il s' #ouvre sur une icône dans une #nouvelle sémose. Aussi le #trait #signifiant, par son détachement même, extériorise une expérience à laquelle il #donne justement accès comme entité #séparée.
N°	χ2	Texte de l'u.c.e. sélectionnée dans la classe 2 ('analyse')
171	26	une #fois les #unités approximativement retenues, on #cherche à étudier la structure des cooccurrences des #mots #pleins dans les #différentes #unités de #contextes. ce #tableau est une #modélisation #possible des trois #aspects de la #répétition, associés à l'objet d' un #discours pour un lecteur: 1) par ce qui est en #lignes;
165	22	2.12 le principe de la #méthode est #simple: le #corpus à analyser est découpé en une #suite de segments de #texte et l'on observe la distribution des #mots #pleins dans ces segments, d' où le nom de la #méthode;
185	21	cet #aspect différentiel de l' #analyse #statistique est, pour nous, #fondamental au #niveau sémiotique, car c'est lui qui #permet de #saisir dans la coupure même des #classes la #stabilisation d un écart.
1	17	L' #objectif que nous nous #proposons est d' aborder la notion de signe de manière très générale en partant de la notion de #répétition, afin de suggérer que l' #approche du sens des #discours dans certains de ses #aspects ne présuppose pas nécessairement l'utilisation de #modèles d' analyse de la #signification linguistique.

Max Reinert (né en 1944) est ingénieur au C.N.R.S. Il est docteur en mathématique option statistique et le créateur de la méthode d'analyse statistique de discours 'ALCESTE' <max.reinert@printemps.uvsq.fr>. Il poursuit ses travaux dans une équipe d'analyse de discours en sociologie du laboratoire PRINTEMPS à l'Université de Versailles Saint-Quentin-en-Yvelines (C.N.R.S.-UPRESA 8085).